



# The Biomedical Ontology Applications (BOA) Framework

Bruno Tavares, Hugo Bastos, Daniel Faria, João Ferreira,  
Tiago Grego, Catia Pesquita, Francisco Couto

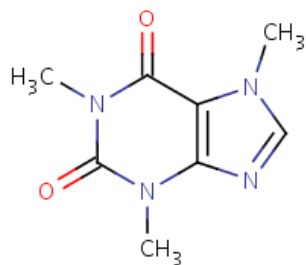
Universidade de Lisboa

July 29, 2011 at ICBO Buffalo USA

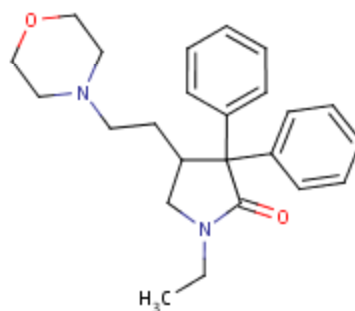
# Semantic Similarity

- Solve the problems of structural similarity
  - Two entities that look similar
    - **Do not imply**
    - They have the same meaning
  - And two entities with similar meaning
    - **Do not imply**
    - that they look similar
- When we want to find similar entities
  - Based on their meaning
  - **Not on** how they look like

# Example: Similar Semantics



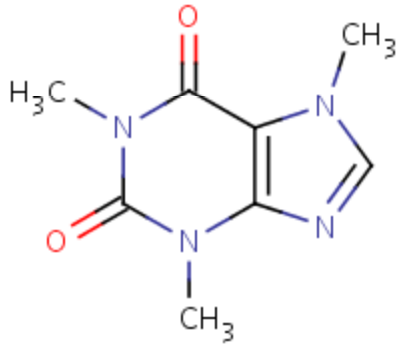
Caffeine (CHEBI:27732)



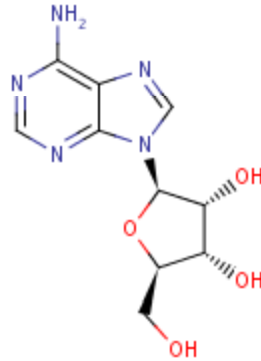
Doxapram (CHEBI:681849)

- Different structure but
- both central nervous system stimulants (CHEBI:35337)

# Example: Similar Structure



caffeine (CHEBI:27732)

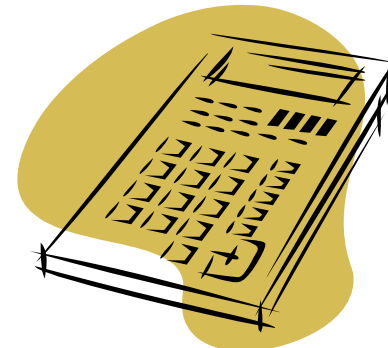


adenosine (CHEBI:16335)

- Similar structure but
- different roles
  - adenosine is an anti-arrhythmia drug (CHEBI:38070)
  - nucleoside (CHEBI:18254)

# Semantic similarity measures

- Input:
  - two ontology concepts
  - or two sets of terms annotating two entities
- Output:
  - a numerical value reflecting the closeness in meaning between them



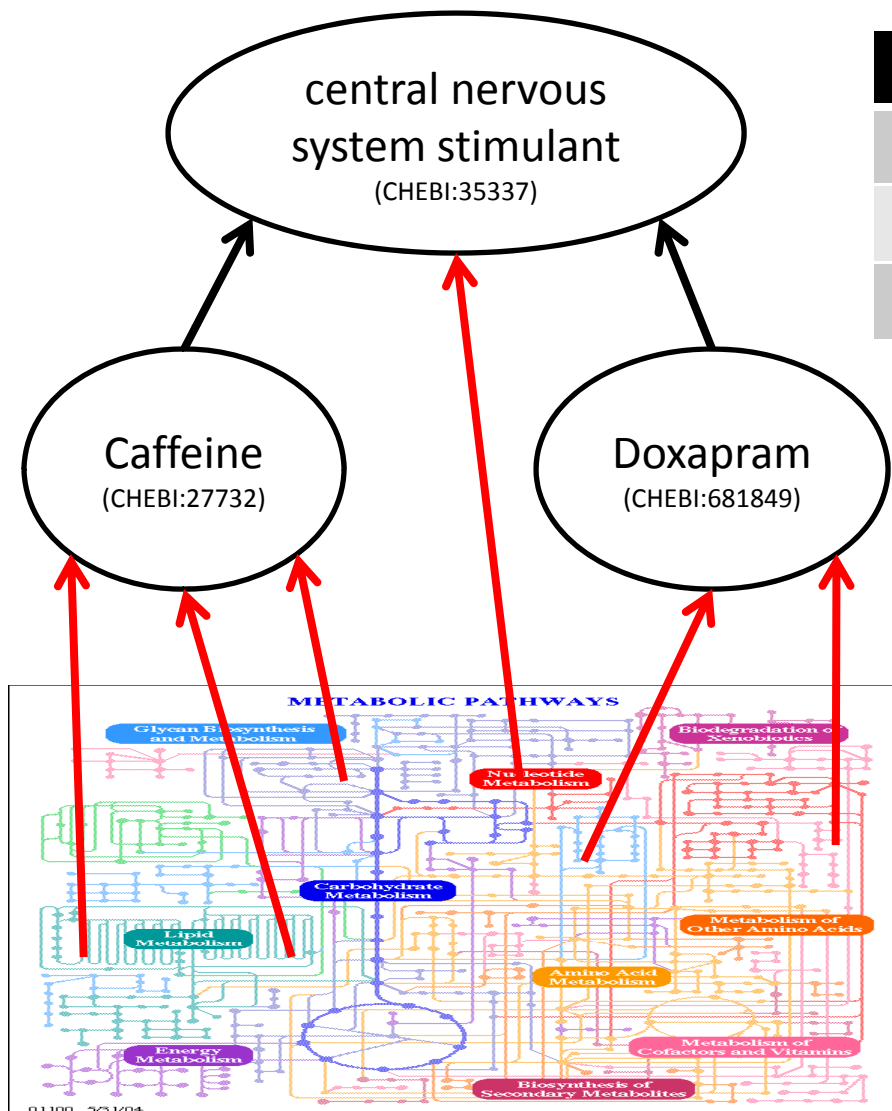
# Information Content

- measures how specific and informative a concept is

$$IC(c) = -\log\left(\frac{freq(c)}{maxFreq}\right)$$

- Inversely proportional to frequency in a given corpus
- The frequency is also propagated to its ancestors
  - IC proportional to the depth of a concept
- Extrinsic IC
  - number of entities mapped to each concept
- Intrinsic IC
  - number of children

# Example



Concept	freq	hfreq	IC
stimulant	1	6	0.0
Caffeine	3	3	0.3
Doxapram	2	2	0.5

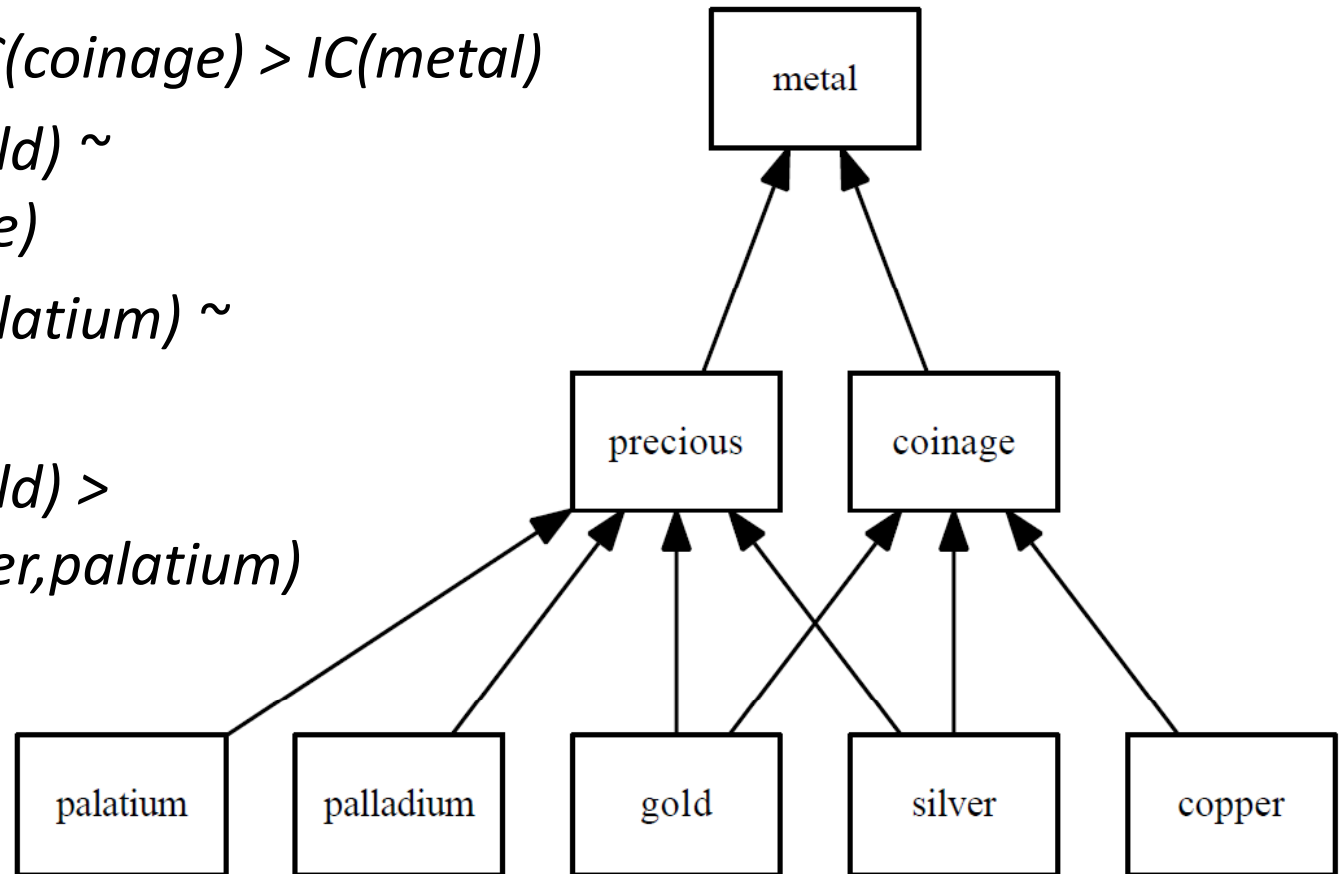
# Similarity based on IC

- Similarity proportional to the
  - IC of the most informative common ancestor (MICA)
    - Shared information between two concepts
    - Resnik
  - Weighted Jaccard index of two sets of concepts
    - Shared information between two entities
    - SimGIC



# Example

- $IC(copper) > IC(coinage) > IC(metal)$
- $Sim(copper, gold) \sim IC(coinage)$
- $Sim(copper, palatium) \sim IC(metal)$
- $Sim(copper, gold) > Sim(copper, palatium)$



## navigation

- [Main Page](#)
- [Publications](#)
- [Members](#)
- [Software](#)
- [Contacts](#)

## information management

- [Research Line](#)
- [CIDS](#)
- [Grease](#)
- [PATI](#)
- [REACTION](#)
- [SIntelliGIS](#)
- [VisMobile](#)

## biomedical informatics

- [Research Line](#)
- [AEB](#)
- [CoherenZy](#)
- [HCM](#)
- [ThermInfo](#)
- [FiBRE](#)
- [EPIWork](#)
- [SABO](#)

## internal information

- [XLDB](#)

## search

xldb.fc.ul.pt/wiki/File:Boa.jpg

[page](#)
[discussion](#)
[view source](#)
[hist](#)

<http://xldb.di.fc.ul.pt/wiki/BOA>

## BOA



Image:Boa.jpg

BOA (Biomedical Ontology Applications) (Aplicações de Ontologias Biomédicas) aims at researching and developing applications to effectively explore all the information contained in Biomedical Ontologies.

## Tutorials

- [Demonstration Video](#)
- [Demo at ICBO 2011](#)

## Web Tools

- [ProteinOn](#) can be used to find interacting proteins, find assigned GO terms and calculate the functional semantic similarity of proteins and to get the information content and calculate the functional semantic similarity of GO terms
- [CMPSim](#) provides a functional similarity measure between chemical compounds and metabolic pathways using ChEBI-based semantic similarity measures
- [GRYFUN](#) for the visualization, filtering and analysis of GO functional annotation profiles of a given protein family
- [CESSM](#) provides a tool for the automated evaluation of GO-based semantic similarity measures

## BOA API Documentation

[BOA API Documentation](#)

## biomedical informatics

- [Research Line](#)
- [AEB](#)
- [CoherenZy](#)
- [HCM](#)
- [ThermInfo](#)
- [FiBRE](#)
- [EPIWork](#)
- [SABO](#)

## internal information

- [XLDB](#)

## search

Go

Search

## toolbox

- [What links here](#)
- [Related changes](#)





## Tutorials

---

- [Demonstration Video](#) 
- [Demo at ICBO 2011](#)

## Web Tools

---

- [ProteinOn](#)  can be used to find interacting proteins, find assigned GO terms and calculate the functional semantic similarity of proteins and to get the information content and calculate the functional semantic similarity of GO terms
- [CMPSim](#)  provides a functional similarity measure between chemical compounds and metabolic pathways using ChEBI-based semantic similarity measures
- [GRYFUN](#)  for the visualization, filtering and analysis of GO functional annotation profiles of a given protein family
- [CESSM](#)  provides a tool for the automated evaluation of GO-based semantic similarity measures

## BOA API Documentation

---

Select protein set to explore:

CAZy

Placeholder 1...

Placeholder 2...

Input set...

**GRYFUN**, **G**Raph **a**nalYzer of **F**UNCTIONal annotation is a webtool that allows the visualization, filtering and subsequent analysis of Gene Ontology (GO) annotation profiles. A GO annotation functional profile consists of the collection of GO terms annotated to proteins in a given protein family (or set).

GO comprehends three orthogonal ontologies - biological process, molecular function and cellular component. These three orthogonal ontology branches of GO are structured as Directed Acyclic Graphs (DAGs).

Hence, GRYFUN's central visualization mode consists of generating and displaying the sub-graphs that subsume all the GO annotations within a given set of proteins for each of these branches. The represented DAGs are comprised of nodes and edges and whereas each node represents a GO term, the connecting directed edges represent taxonomic relationships between them. Relevant GO-based annotation and statistic metrics are also displayed with the graphs.

Thus, these GRYFUN visualizations allow, at quick glance, to grasp both the functional broadness and depth of GO annotation within a given protein set. Moreover, this webtool enables the user to create subsets of proteins associated to terms by clicking their respective nodes on the DAG.

#### Available data:

Currently, only the public resources of the [CAZy](#) database are available for exploration within this tool. Other databases/datasets will be added in the future, and currently under implementation is the input of user sets functionality.

#### WARNING!

This webtool is currently still at Alpha developing stage. Some of the planned features are still under implementation/testing and there are several known bugs, and certainly also some unknown bugs.

All comments, feature suggestions or bug reports are welcomed. [Contact me](#)

Home

About

Sources

Contacts

Select protein set to explore:

CAZy

Placeholder 1...

Placeholder 2...

Input set...

**GRYFUN**, **G**Raph ana**l**yer of **F**UNctional annotation is a webtool that allows the visualization, filtering and subsequent analysis of Gene Ontology (GO) annotation profiles. A GO annotation functional profile consists of the collection of GO terms annotated to proteins in a given protein family (or set).

GO comprehends three orthogonal ontologies - biological process, molecular function and cellular component. These three orthogonal ontology branches of GO are structured as Directed Acyclic

### Enzyme Classes

[Glycoside Hydrolases](#)[GlycosylTransferases](#)[Polysaccharide Lyases](#)[Carbohydrate Esterases](#)

Choose a CAZyme enzyme class from the dropdown menu

use is a resource that supplies family descriptions of structurally- and carbohydrate-binding modules (or functional domains) of grade, modify, or create glycosidic bonds.

This section of GRYFUN allows users to browse through the four Enzyme Classes currently covered by CAZy families of modules that catalyze the breakdown, biosynthesis or modification of carbohydrates and glycoconjugates:

- [Glycoside Hydrolases](#) (GHs) : hydrolysis and/or rearrangement of glycosidic bonds
- [GlycosylTransferases](#) (GTs) : formation of glycosidic bonds
- [Polysaccharide Lyases](#) (PLs) : non-hydrolytic cleavage of glycosidic bonds
- [Carbohydrate Esterases](#) (CEs) : hydrolysis of carbohydrate esters

[Home](#)
[About](#)
[Sources](#)
[Contacts](#)

## Enzyme Classes

[Glycoside Hydrolases](#)
[GlycosylTransferases](#)
[Polysaccharide Lyases](#)
[Carbohydrate Esterases](#)
[PL1](#)
[PL2](#)
[PL3](#)
[PL4](#)
[PL5](#)
[PL6](#)
[PL7](#)
[PL8](#)
[PL9](#)
[PL10](#)
[PL11](#)
[PL12](#)
[PL13](#)
[PL14](#)
[PL15](#)
[PL16](#)
[PL17](#)
[PL18](#)
[PL20](#)
[PL21](#)
[PL22](#)

**base is a resource that supplies family descriptions of  
ed catalytic and carbohydrate-binding modules (or functional  
ymes that degrade, modify, or create glycosidic bonds.**

This section of GRYFUN allows users to browse through the four Enzyme Classes currently covered by CAZy families of modules that catalyze the breakdown, biosynthesis or modification of carbohydrates and glycoconjugates:

[Home](#)

[About](#)

[Sources](#)

[Contacts](#)

**Enzyme Classes**

PL1	PL2	PL3	PL4	PL5	PL6	PL7	PL8	PL9	PL10	PL11
PL12	PL13	PL14	PL15	PL16	PL17	PL18	PL20	PL21	PL22	

**Biological Process**

**Molecular Function**

**Cellular Component**

Generate 'PL1' biological process graph



Home

About

Sources

Contacts

Enzyme Classes

PL1 PL2 PL3 PL4 PL5 PL6 PL7 PL8 PL9 PL10 PL11 PL12 PL13 PL14 PL15 PL16 PL17 PL18  
PL20 PL21 PL22

Biological Process

Molecular Function

Cellular Component

## Family/set: [PL1 \(CAZy\)](#)

**Total family/set size:** 672 protein entries;

**UniProt coverage:** 73.1%; (491 entries);

**Ontology coverage:** 6.99%; (47 annotated entries);

**GOscore:** 0.197 **GOocc:** 0.362

**Root term:** biological\_process

**Root's frequency:** 100% (47)

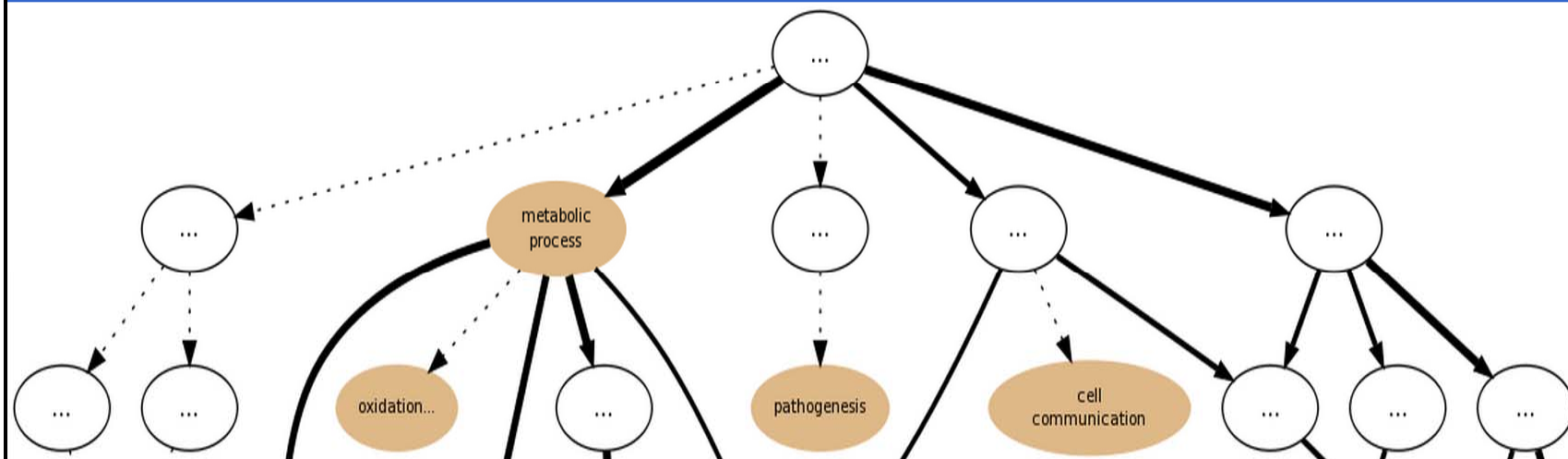
**Bacteria :** ( 27 ) **Eukaryota :** ( 19 ) **Other :** ( 1 )

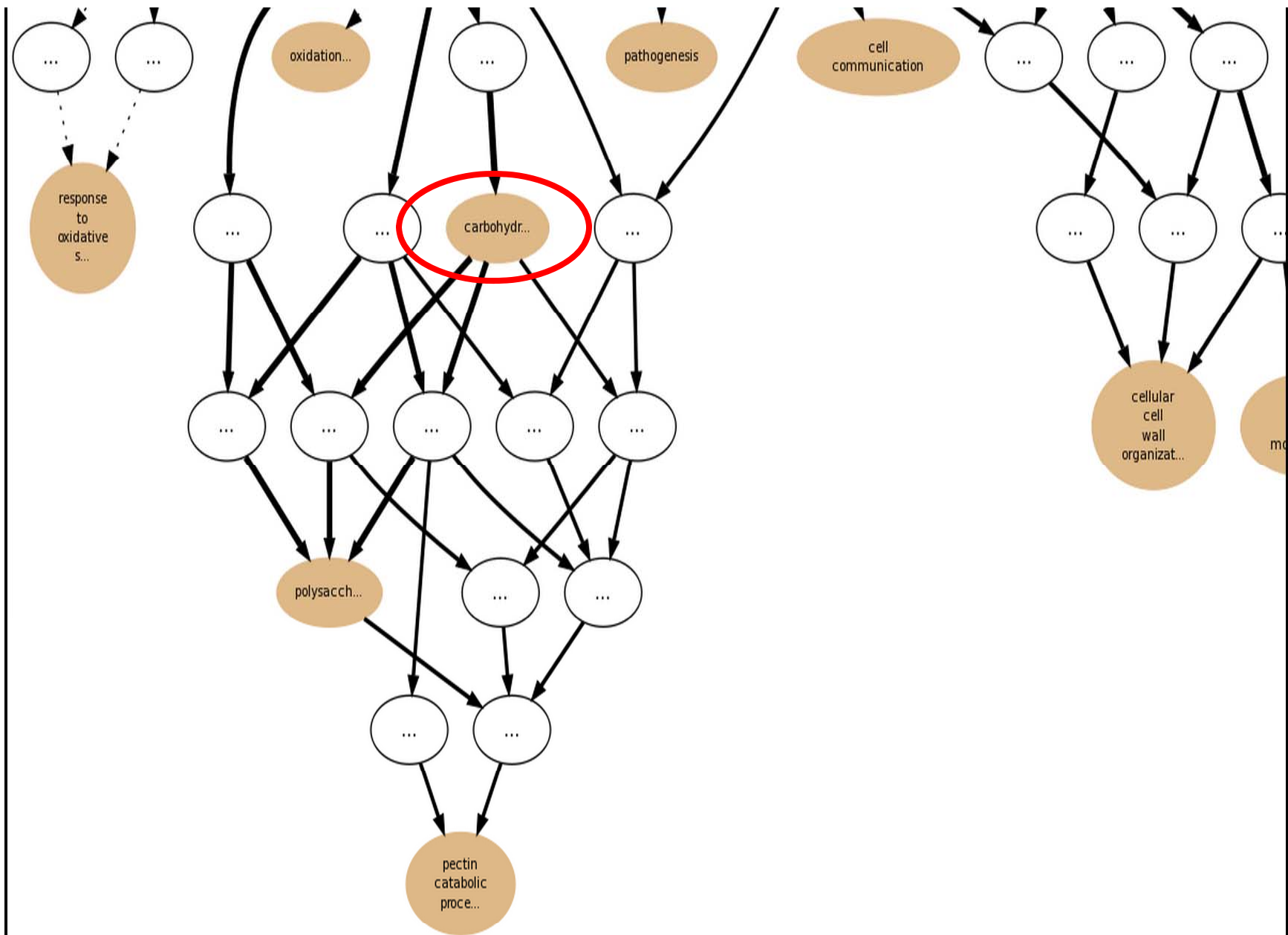
Export  
Graph

Export  
.dot

Fit to  
Width

Original  
Width





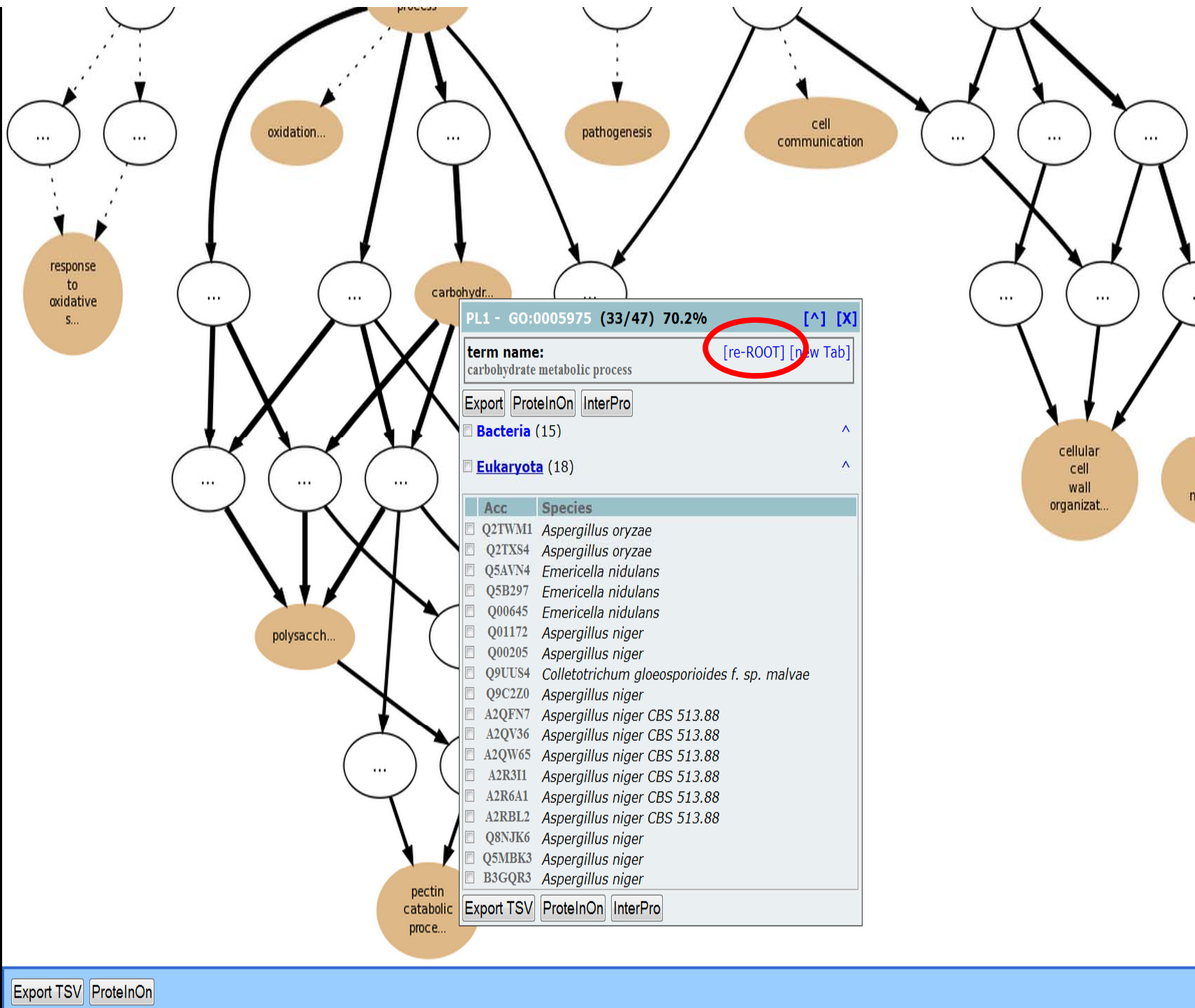
Export TSV ProteInOn

☐ term names

☐ polysaccharide catabolic process

☐ cell wall organization

occ	score
27	0.197
28	0.188



## Family/set: [PL1 \(CAZy\)](#)

**Total family/set size:** 672 protein entries;

**UniProt coverage:** 73.1%; (491 entries);

**Ontology coverage:** 6.99%; (47 annotated entries);

**GOscore:** 0.281 **GOocc:** 0.576

**Root term:** carbohydrate metabolic process

**Root's frequency:** 70.2% (33)

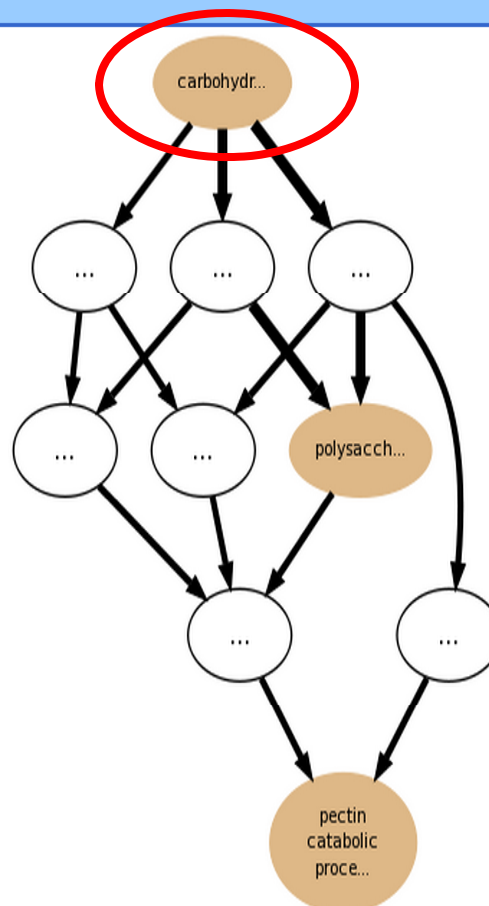
**Bacteria :** ( 15 ) **Eukaryota :** ( 18 )

Export  
Graph

Export  
.dot

Fit to  
Width

Original  
Width



Export TSV

ProteinOn

☐ term names

☐ polysaccharide catabolic process

☐ pectin catabolic process

☐ pectin metabolic process

	occ	score
--	-----	-------

	27	0.281
--	----	-------

	14	0.247
--	----	-------

	14	0.246
--	----	-------

carbohydr...

PL1 - GO:0005975 (33/33) 100% [^] [X]

term name: carbohydrate metabolic process [re-ROOT] [new Tab]

Export ProtInOn InterPro

☒ Bacteria (15) ^

Acc	Species
<input type="checkbox"/> C6WMZ9	Actinosynnema mirum DSM 43827
<input type="checkbox"/> D9SZ13	Micromonospora aurantiaca ATCC 27029
<input type="checkbox"/> D9STE7	Clostridium cellulovorans 743B
<input type="checkbox"/> D9SX08	Clostridium cellulovorans 743B
<input type="checkbox"/> B8I4T1	Clostridium cellulolyticum H10
<input type="checkbox"/> A3DHF2	Clostridium thermocellum ATCC 27405
<input type="checkbox"/> A3DJL9	Clostridium thermocellum ATCC 27405
<input type="checkbox"/> C5BT62	Teredinibacter turnerae T7901
<input type="checkbox"/> A9B3W1	Herpetosiphon aurantiacus ATCC 23779
<input type="checkbox"/> Q21IA8	Saccharophagus degradans 2-40
<input type="checkbox"/> D4LB72	Ruminococcus sp. 18P13
<input type="checkbox"/> D4LCH8	Ruminococcus sp. 18P13
<input type="checkbox"/> D4LD20	Ruminococcus sp. 18P13
<input type="checkbox"/> D4LD22	Ruminococcus sp. 18P13
<input type="checkbox"/> D4LDF6	Ruminococcus sp. 18P13

☒ Eukaryota (18) ^

Acc	Species
<input type="checkbox"/> Q2TWM1	Aspergillus oryzae
<input type="checkbox"/> Q2TXS4	Aspergillus oryzae
<input type="checkbox"/> Q5AVN4	Emericella nidulans
<input type="checkbox"/> Q5B297	Emericella nidulans
<input type="checkbox"/> Q00645	Emericella nidulans
<input type="checkbox"/> Q01172	Aspergillus niger
<input type="checkbox"/> Q00205	Aspergillus niger
<input type="checkbox"/> Q9UUS4	Colletotrichum gloeosporioides f. sp. malvae
<input type="checkbox"/> Q9C2Z0	Aspergillus niger
<input type="checkbox"/> A2QFN7	Aspergillus niger CBS 513.88
<input type="checkbox"/> A2QV36	Aspergillus niger CBS 513.88
<input type="checkbox"/> A2QW65	Aspergillus niger CBS 513.88
<input type="checkbox"/> A2R3I1	Aspergillus niger CBS 513.88
<input type="checkbox"/> A2R6A1	Aspergillus niger CBS 513.88
<input type="checkbox"/> A2RBL2	Aspergillus niger CBS 513.88
<input type="checkbox"/> Q8NJK6	Aspergillus niger
<input type="checkbox"/> Q5MBK3	Aspergillus niger
<input type="checkbox"/> B3GQR3	Aspergillus niger

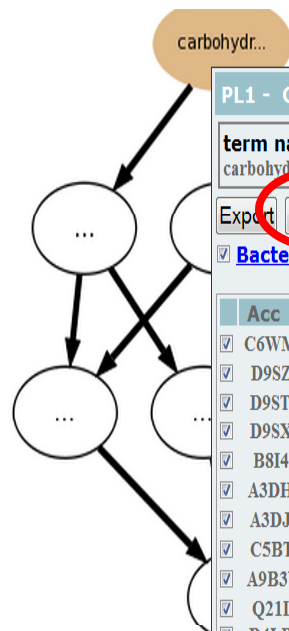
Export TSV ProtInOn InterPro

Export TSV ProtInOn

☒ term names

- ☐ polysaccharide catabolic process
- ☐ pectin catabolic process
- ☐ pectin metabolic process
- ☐ carbohydrate catabolic process
- ☐ cellular polysaccharide catabolic process
- ☐ polysaccharide metabolic process
- ☐ carbohydrate metabolic process
- ☐ cellular polysaccharide metabolic process
- ☐ cellular carbohydrate catabolic process
- ☐ cellular carbohydrate metabolic process

occ	score
27	0.281
14	0.247
14	0.246
27	0.217
14	0.212
27	0.201
33	0.168
14	0.129
14	0.123
14	0.091



PL1 - GO:0005975 (33/33) 100% [^] [X]

term name: carbohydrate metabolic process [re-ROOT] [new Tab]

Export ProtInOn InterPro

☒ Bacteria (15)

Acc	Species
<input checked="" type="checkbox"/> C6WMZ9	Actinosynnema mirum DSM 43827
<input checked="" type="checkbox"/> D9SZ13	Micromonospora aurantiaca ATCC 27029
<input checked="" type="checkbox"/> D9STE7	Clostridium cellulovorans 743B
<input checked="" type="checkbox"/> D9SX08	Clostridium cellulovorans 743B
<input checked="" type="checkbox"/> B8I4T1	Clostridium cellulolyticum H10
<input checked="" type="checkbox"/> A3DHF2	Clostridium thermocellum ATCC 27405
<input checked="" type="checkbox"/> A3DJL9	Clostridium thermocellum ATCC 27405
<input checked="" type="checkbox"/> C5BT62	Teredinibacter turnerae T7901
<input checked="" type="checkbox"/> A9B3W1	Herpetosiphon aurantiacus ATCC 23779
<input checked="" type="checkbox"/> Q21IA8	Saccharophagus degradans 2-40
<input checked="" type="checkbox"/> D4LB72	Ruminococcus sp. 18P13
<input checked="" type="checkbox"/> D4LCH8	Ruminococcus sp. 18P13
<input checked="" type="checkbox"/> D4LD20	Ruminococcus sp. 18P13
<input checked="" type="checkbox"/> D4LD22	Ruminococcus sp. 18P13
<input checked="" type="checkbox"/> D4LDF6	Ruminococcus sp. 18P13

☒ Eukaryota (18)

Acc	Species
<input checked="" type="checkbox"/> Q2TWM1	Aspergillus oryzae
<input checked="" type="checkbox"/> Q2TXS4	Aspergillus oryzae
<input checked="" type="checkbox"/> Q5AVN4	Emericella nidulans
<input checked="" type="checkbox"/> Q5B297	Emericella nidulans
<input checked="" type="checkbox"/> Q00645	Emericella nidulans
<input checked="" type="checkbox"/> Q01172	Aspergillus niger
<input checked="" type="checkbox"/> Q00205	Aspergillus niger
<input checked="" type="checkbox"/> Q9UUS4	Colletotrichum gloeosporioides f. sp. malvae
<input checked="" type="checkbox"/> Q9C2Z0	Aspergillus niger
<input checked="" type="checkbox"/> A2QFN7	Aspergillus niger CBS 513.88
<input checked="" type="checkbox"/> A2QV36	Aspergillus niger CBS 513.88
<input checked="" type="checkbox"/> A2QW65	Aspergillus niger CBS 513.88
<input checked="" type="checkbox"/> A2R3I1	Aspergillus niger CBS 513.88
<input checked="" type="checkbox"/> A2R6A1	Aspergillus niger CBS 513.88
<input checked="" type="checkbox"/> A2RBL2	Aspergillus niger CBS 513.88
<input checked="" type="checkbox"/> Q8NJK6	Aspergillus niger
<input checked="" type="checkbox"/> Q5MBK3	Aspergillus niger
<input checked="" type="checkbox"/> B3GQR3	Aspergillus niger

Export TSV ProtInOn InterPro

Export TSV ProtInOn

☐ term names

- ☐ polysaccharide catabolic process
- ☐ pectin catabolic process
- ☐ pectin metabolic process
- ☐ carbohydrate catabolic process
- ☐ cellular polysaccharide catabolic process
- ☐ polysaccharide metabolic process
- ☐ carbohydrate metabolic process
- ☐ cellular polysaccharide metabolic process
- ☐ cellular carbohydrate catabolic process
- ☐ cellular carbohydrate metabolic process

occ	score
27	0.281
14	0.247
14	0.246
27	0.217
14	0.212
27	0.201
33	0.168
14	0.129
14	0.123
14	0.091

Home

About

Sources

Team

### Step 1: Query

compute protein semantic similarity ▾

### Step 2: Options

Measure: simGIC ▾

GO type: Molecular Function ▾

☐ Ignore IEA

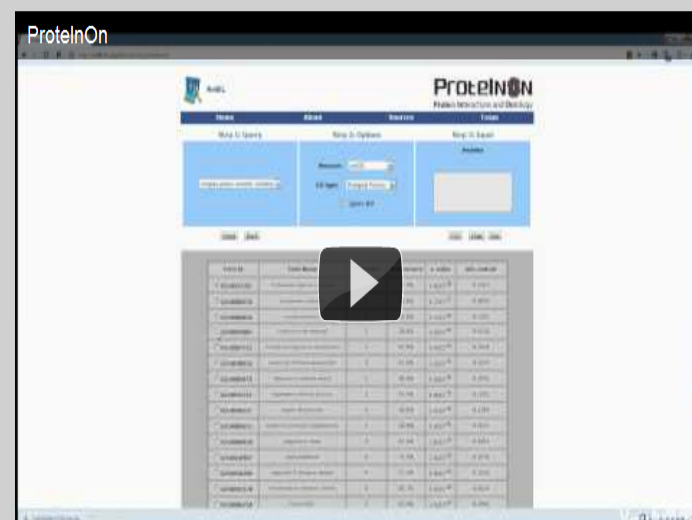
### Step 3: Input

#### Proteins

C6WMZ9, D9SZ13, D9STE7, D9SX  
08, B8I4T1, A3DHF2, A3DJL9, C  
5BT62, A9B3W1, Q21IA8, D4LB7  
2, D4LCH8, D4LD20, D4LD22, D4  
LDF6, Q2TWM1, Q2TXS4, Q5AVN4

Run

The query **compute protein semantic similarity** returns the semantic similarity scores between all proteins entered, in matrix format.



## Step 1: Query

## Step 2: Options

Measure: GO type: ☐ Ignore IEA

## Step 3: Input

## Proteins

Reset

Back

XML

TSV

Run

☐ Select All

Protein 1	Protein 2	Score
<<prev 1/53 next>>		
<input type="checkbox"/> C6WMZ9	<input type="checkbox"/> D9SZ13	100%
<input type="checkbox"/> C6WMZ9	<input type="checkbox"/> A9B3W1	100%
<input type="checkbox"/> B8I4T1	<input type="checkbox"/> D4LB72	100%
<input type="checkbox"/> B8I4T1	<input type="checkbox"/> D4LCH8	100%
<input type="checkbox"/> B8I4T1	<input type="checkbox"/> D4LDF6	100%
<input type="checkbox"/> A3DHF2	<input type="checkbox"/> A3DJL9	100%
<input type="checkbox"/> A3DHF2	<input type="checkbox"/> C5BT62	100%
<input type="checkbox"/> A3DJL9	<input type="checkbox"/> C5BT62	100%



## Step 1: Query

find GO terms representativity  
GO terms to ChEBI  
compute term semantic similarity  
compute protein semantic similarity  
find interacting proteins  
find GO terms representativity

Reset

Back

## Step 2: Options

Measure: simGIC

GO type: Molecular Function

☐ Ignore IEA

## Step 3: Input

## Proteins

XML

TSV

Run

☒ Select All

Protein 1	Protein 2	Score
<<prev 1/53 next>>		
<input checked="" type="checkbox"/> C6WMZ9	<input checked="" type="checkbox"/> D9SZ13	100%
<input checked="" type="checkbox"/> C6WMZ9	<input checked="" type="checkbox"/> A9B3W1	100%
<input checked="" type="checkbox"/> B8I4T1	<input checked="" type="checkbox"/> D4LB72	100%
<input checked="" type="checkbox"/> B8I4T1	<input checked="" type="checkbox"/> D4LCH8	100%
<input checked="" type="checkbox"/> B8I4T1	<input checked="" type="checkbox"/> D4LDF6	100%
<input checked="" type="checkbox"/> A3DHF2	<input checked="" type="checkbox"/> A3DJL9	100%
<input checked="" type="checkbox"/> A3DHF2	<input checked="" type="checkbox"/> C5BT62	100%
<input checked="" type="checkbox"/> A3DJL9	<input checked="" type="checkbox"/> C5BT62	100%

## Step 1: Query

## Step 2: Options

**Measure:** 
**GO type:** 
☐ Ignore IEA

## Step 3: Input

## Proteins

Reset

Back

TSV

Chart

Run

☐ Select All

Term Id	Term Name	# Proteins	%Occurence	e-value	info content
<<prev           1/1           next>>					
<input type="checkbox"/> GO:0047490	pectin lyase activity	12	36.4%	0.0x10 <sup>+00</sup>	0.7461
<input type="checkbox"/> GO:0016837	carbon-oxygen lyase activity, acting on polysaccharides	17	51.5%	0.0x10 <sup>+00</sup>	0.5113
<input type="checkbox"/> GO:0016829	lyase activity	27	81.8%	0.0x10 <sup>+00</sup>	0.2020
<input type="checkbox"/> GO:0030570	pectate lyase activity	5	15.2%	8.9x10 <sup>-16</sup>	0.5929
<input type="checkbox"/> GO:0004553	hydrolase activity, hydrolyzing O-glycosyl compounds	12	36.4%	1.5x10 <sup>-14</sup>	0.2689
<input type="checkbox"/> GO:0030246	carbohydrate binding	9	27.3%	2.3x10 <sup>-12</sup>	0.3087
<input type="checkbox"/> GO:0030247	polysaccharide binding	5	15.2%	2.5x10 <sup>-10</sup>	0.4340
<input type="checkbox"/> GO:0003824	catalytic activity	33	100.0%	1.1x10 <sup>-06</sup>	0.0262

[Home](#)

[About](#)

[Sources](#)

[Contacts](#)

**Enzyme Classes**

PL1	PL2	PL3	PL4	PL5	PL6	PL7	PL8	PL9	PL10	PL11
PL12	PL13	PL14	PL15	PL16	PL17	PL18	PL20	PL21	PL22	

[Biological Process](#)

[Molecular Function](#)

[Cellular Component](#)

Generate 'PL1' molecular function graph

Enzyme Classes

PL1 PL2 PL3 PL4 PL5 PL6 PL7 PL8 PL9 PL10 PL11 PL12 PL13 PL14 PL15 PL16 PL17 PL18  
PL20 PL21 PL22

Biological Process

Molecular Function

Cellular Component

Family/set: **PL1 (CAZy)**

Total family/set size: 672 protein entries;

UniProt coverage: 73.1%; (491 entries);

Ontology coverage: 58.2%; (391 annotated entries);

GOscore: 0.196 GOocc: 0.146

Root term: molecular\_function

Root's frequency: 100% (391)

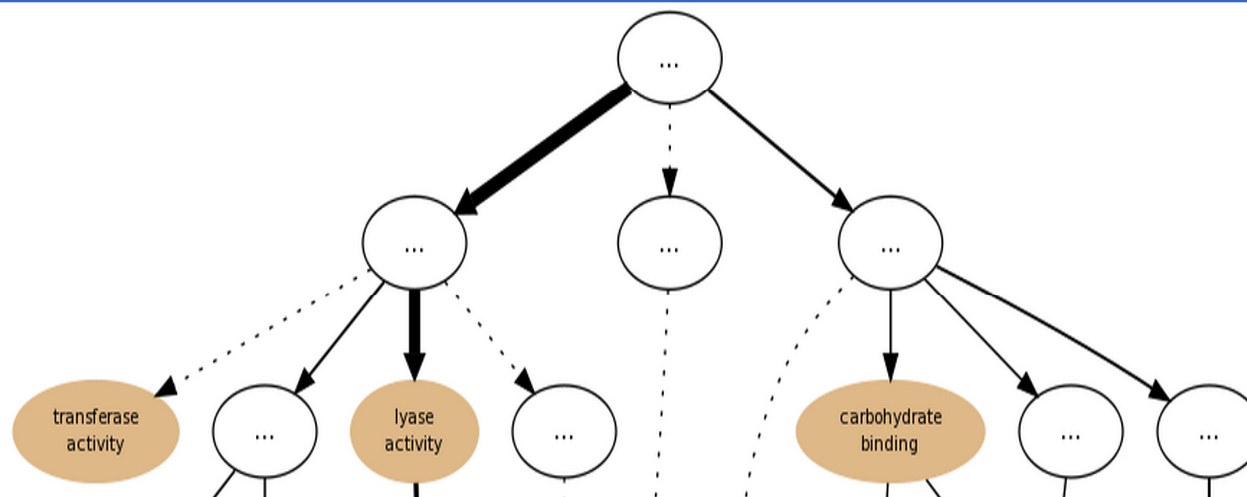
Archaea : ( 2 ) Bacteria : ( 276 ) Eukaryota : ( 112 ) Other : ( 1 )

Export  
Graph

Export  
.dot

Fit to  
Width

Original  
Width





Export TSV ProteinOn

<input checked="" type="checkbox"/> term names	occ	score
<input type="checkbox"/> lyase activity	379	0.196
<input type="checkbox"/> carbon-oxygen lyase activity, acting on polysaccharides	119	0.156
<input type="checkbox"/> pectate lyase activity	97	0.147
<input type="checkbox"/> carbon-oxygen lyase activity	119	0.086
<input type="checkbox"/> pectin lyase activity	23	0.044
<input type="checkbox"/> catalytic activity	389	0.026
<input type="checkbox"/> pectinesterase activity	11	0.014
<input type="checkbox"/> carbohydrate binding	16	0.013
<input type="checkbox"/> carboxylesterase activity	11	0.010
<input type="checkbox"/> hydrolase activity, acting on glycosyl bonds	13	0.009
<input type="checkbox"/> metal ion binding	29	0.008
<input type="checkbox"/> cation binding	29	0.008
<input type="checkbox"/> ion binding	29	0.008
<input type="checkbox"/> hydrolase activity, hydrolyzing O-glycosyl compounds	12	0.008
<input type="checkbox"/> hydrolase activity	24	0.006
<input type="checkbox"/> hydrolase activity, acting on ester bonds	11	0.006
<input type="checkbox"/> polysaccharide binding	5	0.006
<input type="checkbox"/> pattern binding	5	0.006
<input type="checkbox"/> binding	45	0.004
<input type="checkbox"/> sugar binding	3	0.003
<input type="checkbox"/> cellulose binding	1	0.001
<input type="checkbox"/> peroxidase activity	1	0.001
<input type="checkbox"/> oxidoreductase activity, acting on peroxide as acceptor	1	0.001
<input type="checkbox"/> antioxidant activity	1	0.001
<input type="checkbox"/> heme binding	1	0.001
<input type="checkbox"/> tetrapyrrole binding	1	0.001
<input type="checkbox"/> iron ion binding	1	0.000
<input type="checkbox"/> transition metal ion binding	1	0.000
<input type="checkbox"/> oxidoreductase activity	1	0.000
<input type="checkbox"/> transferase activity	1	0.000
<input type="checkbox"/> molecular_function	391	0.000



Export To **ProteinOn**

<input checked="" type="checkbox"/> term names	occ	score
<input checked="" type="checkbox"/> lyase activity	379	0.196
<input checked="" type="checkbox"/> carbon-oxygen lyase activity, acting on polysaccharides	119	0.156
<input checked="" type="checkbox"/> pectate lyase activity	97	0.147
<input checked="" type="checkbox"/> carbon-oxygen lyase activity	119	0.086
<input checked="" type="checkbox"/> pectin lyase activity	23	0.044
<input checked="" type="checkbox"/> catalytic activity	389	0.026
<input checked="" type="checkbox"/> pectinesterase activity	11	0.014
<input checked="" type="checkbox"/> carbohydrate binding	16	0.013
<input checked="" type="checkbox"/> carboxylesterase activity	11	0.010
<input checked="" type="checkbox"/> hydrolase activity, acting on glycosyl bonds	13	0.009
<input checked="" type="checkbox"/> metal ion binding	29	0.008
<input checked="" type="checkbox"/> cation binding	29	0.008
<input checked="" type="checkbox"/> ion binding	29	0.008
<input checked="" type="checkbox"/> hydrolase activity, hydrolyzing O-glycosyl compounds	12	0.008
<input checked="" type="checkbox"/> hydrolase activity	24	0.006
<input checked="" type="checkbox"/> hydrolase activity, acting on ester bonds	11	0.006
<input checked="" type="checkbox"/> polysaccharide binding	5	0.006
<input checked="" type="checkbox"/> pattern binding	5	0.006
<input checked="" type="checkbox"/> binding	45	0.004
<input checked="" type="checkbox"/> sugar binding	3	0.003
<input checked="" type="checkbox"/> cellulose binding	1	0.001
<input checked="" type="checkbox"/> peroxidase activity	1	0.001
<input checked="" type="checkbox"/> oxidoreductase activity, acting on peroxide as acceptor	1	0.001
<input checked="" type="checkbox"/> antioxidant activity	1	0.001
<input checked="" type="checkbox"/> heme binding	1	0.001
<input checked="" type="checkbox"/> tetrapyrrole binding	1	0.001
<input checked="" type="checkbox"/> iron ion binding	1	0.000
<input checked="" type="checkbox"/> transition metal ion binding	1	0.000
<input checked="" type="checkbox"/> oxidoreductase activity	1	0.000
<input checked="" type="checkbox"/> transferase activity	1	0.000
<input checked="" type="checkbox"/> molecular_function	391	0.000

Home

About

Sources

Team

### Step 1: Query

GO terms to ChEBI  
GO terms to ChEBI  
compute term semantic similarity  
compute protein semantic similarity  
find interacting proteins  
find GO terms representativity

### Step 2: Options

Measure: simGIC

GO type: Molecular Function

☐ Ignore IEA

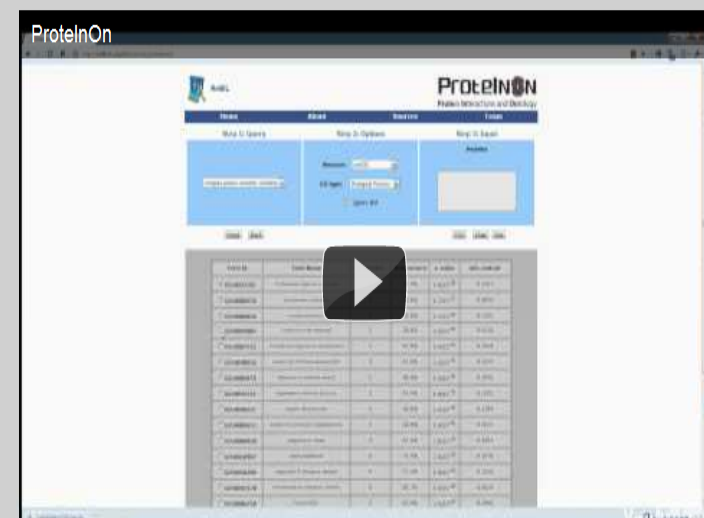
### Step 3: Input

#### Terms

GO:0016829,GO:0016837,GO:  
0030570,GO:0016835,GO:004  
7490,GO:0003824,GO:003059  
9,GO:0030246,GO:0004091,G  
O:0016798,GO:0046872,GO:0

Run

The query **compute term semantic similarity** returns the semantic similarity scores between all GO terms entered, in matrix format.



Home

About

Sources

Team

Step 1: Query

compute protein semantic similarity

Step 2: Options

Measure: simGIC

GO type: Molecular Function

☐ Ignore IEA

Step 3: Input

Terms

GO:0016829,GO:0016837,GO:  
0030570,GO:0016835,GO:004  
7490,GO:0003824,GO:003059  
9,GO:0030246,GO:0004091,G  
O:0016798,GO:0046872,GO:0

Run

[Continue to CMPSim](#)



# CMPSim

Chemical and Metabolic Pathway Similarity

::: alpha version :::

## Compare compounds

30413,26932

Measure:

Run

## Search

Search

Try an [Advanced Search](#).

## Tools

- » [Calculate similarity between compounds](#)
- » [Calculate similarity between pathways](#)
- » [Characterize pathways](#)
- » [Pathway index](#)

This webtool is brought to you by

[XLDB - Lasige](#)

**BOA** BIOMEDICAL  
ONTOLOGY  
APPLICATIONS

# CMPSim




Chemical and Metabolic Pathway Similarity

::: alpha version :::

## Compare compounds

30413,26932

Measure:

Compound 1 	Compound 2 	Score 
<<prev	1/1	next>>
<b>CHEBI:30413</b>	<b>CHEBI:26932</b>	19.71%

## Search

Try an [Advanced Search](#).

## Tools

- » [Calculate similarity between compounds](#)
- » [Calculate similarity between pathways](#)
- » [Characterize pathways](#)
- » [Pathway index](#)

This webtool is brought to you by

[XLDB - Lasige](#)

- ChEBI Home
  - Advanced Search
  - Browse
  - Submissions
  - Downloads
  - Documentation
  - Developer Resources
  - Preferences
  - Contact ChEBI
- Printer Friendly View

EBI > Databases > Small Molecules > ChEBI > Main

heme (CHEBI:30413)

Search ChEBI here!

Search ChEBI

Search for ★★ only

Main

Automatic Xrefs

ChEBI Name

heme

ChEBI ID

CHEBI:30413

Definition

A heme is any tetrapyrrolic chelate of iron.

Stars

★★★ This entity has been manually annotated by the ChEBI Team.

Secondary ChEBI IDs

CHEBI:14386, CHEBI:24491

ChEBI Ontology

Tree view

Outgoing

heme (CHEBI:30413) is a metallotetrapyrrole (CHEBI:33909)

heme (CHEBI:30413) is a metalloporphyrin (CHEBI:25216)

heme (CHEBI:30413) has role prosthetic group (CHEBI:26348)

Incoming

siroheme(6-) (CHEBI:60052) is a heme (CHEBI:30413)

hemoprotein (CHEBI:35137) has part heme (CHEBI:30413)

ferroheme (CHEBI:38573) is a heme (CHEBI:30413)

ferriheme (CHEBI:38574) is a heme (CHEBI:30413)

heme-amino acid (CHEBI:24485) is a heme (CHEBI:30413)

alpha-oxyprotohemin IX (CHEBI:22465) is a heme (CHEBI:30413)

etiohemin I (CHEBI:24008) is a heme (CHEBI:30413)

heme a (CHEBI:24479) is a heme (CHEBI:30413)

heme o (CHEBI:24480) is a heme (CHEBI:30413)

heme-thiolate prosthetic group (CHEBI:36073) is a heme (CHEBI:30413)

heme b (CHEBI:26355) is a heme (CHEBI:30413)

siroheme (CHEBI:28599) is a heme (CHEBI:30413)

IUPAC Names

heme

hemes

Synonyms

haem

hæm

Sources

IUPAC

ChEBI

## biomedical informatics

- Research Line
- AEB
- CoherenZy
- HCM
- ThermlInfo
- FiBRE
- EPIWork
- SABO

## internal information

- XLDB

## search

Go


Search

## toolbox





- What links here
- Related changes

## Tutorials

<http://xldb.di.fc.ul.pt/wiki/BOA>

- [Demonstration Video](#) 
- [Demo at ICBO 2011](#)

## Web Tools

- [ProteinOn](#)  can be used to find interacting proteins, find assigned GO terms and calculate the functional semantic similarity of proteins and to get the information content and calculate the functional semantic similarity of GO terms
- [CMPSim](#)  provides a functional similarity measure between chemical compounds and metabolic pathways using ChEBI-based semantic similarity measures
- [GRYFUN](#)  for the visualization, filtering and analysis of GO functional annotation profiles of a given protein family
- [CESSM](#)  provides a tool for the automated evaluation of GO-based semantic similarity measures

## BOA API Documentation